

L'apprentissage au secours de la réduction de dimension pour des problèmes d'optimisation

Ala Ben Abbes¹, Emmanuel Rachelson², Sébastien Diemer¹

¹ EDF R&D — Optimisation, Simulation, Risques et Statistiques
prenom.nom@edf.fr

² Université de Liège — Institut Montefiore
emmanuel.rachelson@ulg.ac.be

Résumé : Pour assurer la stabilité du réseau électrique, la production doit être ajustée en quasi temps réel à la consommation. Cet ajustement ne peut porter que sur un nombre limité de centrales et doit être effectué dans des délais réduits. La combinatoire du problème rend la recherche d'un optimum économique par des méthodes d'optimisation classiques très difficile. Ce travail cherche à montrer l'intérêt d'utiliser des algorithmes d'apprentissage supervisé performants comme le *Boosting*, pour sélectionner les centrales à redéclarer. Cette sélection préalable permet ensuite de réduire considérablement le temps de l'optimisation des programmes de production tout en garantissant l'optimalité économique.

Mots-clés : Selection de variables, Boosting, Gestion de production.

1 Introduction

Les producteurs d'électricité en France sont tenus de fournir quotidiennement leurs programmes de production la veille pour le lendemain. Cette production prévisionnelle doit par ailleurs être égale à la demande des clients à chaque instant de la journée. L'apparition d'un certain nombre d'aléas en cours de journée remet néanmoins en cause cet équilibre entre production et demande. Il devient alors indispensable de replanifier le programme de production de certaines unités. Cependant, pour garantir la stabilité du réseau, le nombre des unités redéclarées par jour est limité. Ce problème de replanification, dit "infra-journalier", est particulièrement complexe à résoudre de façon optimale pour le producteur. Il s'agit en effet de trouver un nouveau programme de production optimal (*i.e.* qui minimise le coût total de la production) tout en respectant cette nouvelle contrainte portant sur le nombre de re-déclarations. En terme de complexité, il s'agit d'un problème linéaire mixte à 1.000.000 de variables et 2.000.000 de contraintes et cette complexité est incompatible avec les fortes contraintes opérationnelles sur le temps de calcul. L'idée que nous développons ici consiste à tenter de prédire quelles seront les unités participant à la réorganisation du parc électrique face à la nouvelle demande. Armés de cette prédiction, nous pouvons diminuer le temps de calcul nécessaire à une replanification en ne résolvant qu'un problème en dimension réduite.

2 Méthode et outils de résolution

Le principe de notre démarche. Le principe que nous développons ici consiste à capitaliser de l'expérience sur plusieurs résolutions successives du problème de re-planification *hors ligne* afin de construire un prédicteur qui permettra alors *en ligne* de sélectionner les variables pertinentes pour l'optimisation. La première étape de cette démarche est illustrée par la figure 1 et consiste à construire classiquement un plan de production pour la journée du lendemain sur la base d'une demande prévisionnelle de référence, puis à considérer un certain nombre d'aléas portant sur cette demande. On résout alors le problème infra-journalier pour chacun de ces aléas, ce qui fournit autant de plans de production corrigés. Notons que ces traitements se font hors ligne.

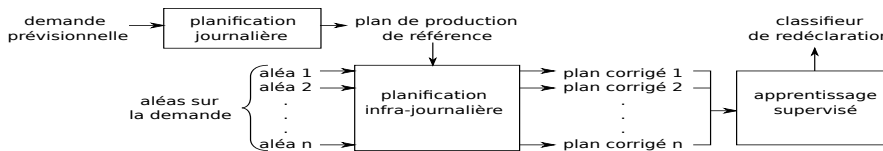


FIG. 1 – Construction hors ligne du sélecteur d'Entité D'Ajustement

On obtient ainsi des *plans de redéclaration* correspondant à chacun des aléas considérés. Chacun de ces plans respecte la contrainte du nombre maximum de redéclarations. On peut ainsi en extraire une information binaire pour chaque *entité d'ajustement* (EDA) indiquant si l'EDA a été ou non redéclarée en présence de l'aléa en question. On construit ainsi une base d'apprentissage globale où chaque exemple décrit une EDA face à un aléa. Ainsi, on obtient $n_{alea} \times n_{EDA}$ exemples et on associe à chacun le label correspondant à la redéclaration de l'EDA ou non. On entraîne alors un classifieur sur cette base d'apprentissage. Ce classifieur infère donc, pour les caractéristiques d'une EDA et d'un aléa donnés, la participation ou non de l'EDA au planning optimal de redéclaration. Une fois ce *sélecteur d'EDA* obtenu et le plan de référence en cours d'exécution, le sélecteur d'EDA permet d'effectuer une sélection de variables en ligne, pour rendre possible la résolution du problème infra-journalier, comme illustré figure 2. Lorsqu'en cours de journée, la demande observée varie par rapport à la demande de référence, le sélecteur d'EDA utilise la donnée de cet aléa et passe en revue les EDA disponibles pour prédire lesquelles sont susceptibles de participer à la redéclaration optimale. Le problème infra-journalier est alors résolu en n'utilisant que les variables correspondant à ces EDA, entraînant un gain important en complexité.



FIG. 2 – Sélection de variables pour la résolution en ligne

Construction du sélecteur d'EDA. Le schéma de principe du paragraphe précédent illustre le processus hors ligne / en ligne mis en oeuvre afin de résoudre le problème infra-journalier. L'élément crucial de cette architecture est le sélecteur d'EDA ; un soin particulier doit donc être apporté au choix de la méthode d'apprentissage supervisé mise en oeuvre. Le méta-algorithme du Boosting (Schapire, 2003), permettant de rendre meilleure la prédiction de tout autre algorithme d'apprentissage supervisé et de construire ainsi des classifieurs *PAC-forts*¹ (Valiant, 1984) a été choisi pour notre étude. Il repose sur l'hypothèse que l'algorithme de classification sur lequel il est utilisé est *PAC-faible*, c'est-à-dire meilleur qu'un tirage aléatoire. Le principe du Boosting est alors de construire des ensembles pondérés de classifieurs PAC-faibles dont le vote converge vers un classifieur *PAC-fort*. ADABOOST (Freund & Schapire, 1996) est le principal algorithme de Boosting. Il s'agit d'une méthode adaptée à notre problème car il requiert peu de paramètres et permet de bonnes performances en présence de données faiblement bruitées. Le choix des classifieurs faibles utilisés avec ADABOOST a également une influence. A titre expérimental, nous avons utilisé à chaque fois plusieurs classifieurs faibles, en gardant celui qui présentait la meilleure erreur d'apprentissage. Ainsi, à chaque itération d'ADABOOST, un choix est fait entre un arbre de décision, une machine à vecteur supports et un ensemble de règles de décisions "métier" issues de l'expérience acquise en pratique par les exploitants.

Validation. Nous souhaitons qu'un exemple de la base d'apprentissage représente une Entité d'Ajustement face à un aléa quelconque. Les variables correspondant à un exemple décrivent donc l'EDA concernée (puissance maximale et minimale, coûts, production de référence...) et l'aléa (valeur de l'aléa sur l'ensemble de l'horizon considéré). Afin d'assurer la richesse de la base d'apprentissage, nous avons utilisé 120 aléas historiques. L'énergie introduite par ces aléas est significative et les aléas sont suffisamment divers pour représenter la réalité. Notre maquette modélise 27 EDA, on dispose donc de 120×27 exemples pour cette étude numérique. Pour chaque exemple, on détermine la valeur du label binaire (EDA redéclarée ou non) avec la maquette présentée plus haut. Cette valeur permet d'effectuer l'apprentissage et servira à valider l'approche durant la phase de test. La validation est faite de manière croisée (leave-one-out).

3 Expériences sur la maquette "infra-journalier"

L'intérêt de cette approche est d'utiliser le vecteur réduit d'EDA obtenu grâce au sélecteur, afin de ré-optimiser le programme de production pour faire face à l'aléa. Pour évaluer les performances de notre algorithme, il est alors intéressant de ne pas simplement considérer l'erreur d'apprentissage ou de généralisation mais le coût du programme de production obtenu en ne modifiant que les EDA indiquées par le classifieur. Le coût ainsi obtenu est à comparer au coût optimal obtenu par la maquette qui résout le problème exact de façon frontale. La figure 3 présente la distribution de l'écart entre le coût obtenu par notre approche et le coût optimal pour chacun des 120 aléas considérés. En moyenne, les résultats restent très honorables puisqu'on obtient un écart moyen de 1000 euros par rapport à un coût optimal qui se situe autour de 3 millions d'euros.

¹Probablement Approximativement Corrects.

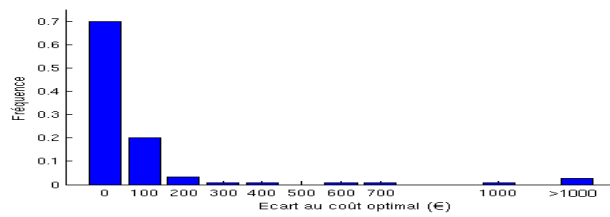


FIG. 3 – Distribution des écarts au coût optimal

Le second objectif de cette approche était de diminuer les temps de calcul importants dûs à la ré-optimisation sur le parc entier. Sur la maquette utilisée, l'optimisation exacte dure 1h, l'approche par Boosting environ 129s (0.24s pour la sélection des EDA et 128.07s pour l'optimisation réduite). Ce gain notable (facteur de 1/30) est très encourageant puisqu'il constitue, rappelons le, la principale motivation de ces travaux.

4 Conclusion

Cet article vise à démontrer dans la pratique l'intérêt des méthodes d'apprentissage supervisé pour la sélection de variables en optimisation (une étude similaire a été menée dans (Cornelusse *et al.*, 2009)). En termes d'optimalité du coût et de temps de calcul, les résultats sont en effet très encourageants : l'écart de coût par rapport à l'optimal est très faible voire nul sur 70% des aléas et le temps de calcul est réduit d'un facteur 30.

Les résultats prometteurs de ce travail ouvrent plusieurs pistes de recherche et d'amélioration et soulèvent des questions pour la suite des travaux : intérêt de la séparation du classifieur général en plusieurs classifieurs (un par EDA), lien avec des méthodes de Boosting pour la classification multi-classes comme ADABOOST.MH ou encore possibilité de réutiliser un classifieur sur plusieurs jours.

Remerciements Emmanuel Rachelson remercie le Réseau de recherches DYSCO, financé par le Inter-university Attraction Poles Programme du bureau scientifique de l'état Belge. Les auteurs remercient l'équipe OSIRIS d'EDF R&D.

Références

- CORNELUSSE B., VIGNAL G., DEFOURNY B. & WEHENKEL L. (2009). Supervised learning of intra-daily recourse strategies for generation management under uncertainties. In *IEEE Power Tech Conference*.
- FREUND Y. & SCHAPIRE R. E. (1996). Experiments with a New Boosting Algorithm. In *International Conference on Machine Learning*.
- SCHAPIRE R. E. (2003). Boosting Approach to Machine Learning : An Overview. *Nonlinear Estimation and Classification*.
- VALIANT L. G. (1984). A Theory of the Learnable. *Communications of the ACM*, **27**(11), 1134–1142.